Benchmarking One-Step Library Preparation Method Against Fragmentation-Ligation Workflows for Genomic Selection Using Low-Pass Whole Genome Sequencing (lpWGS)

Michelle Rahardja, Jack Leonard, Yanyan Liu, Rebecca Feeley, Jenna Couture, Stella Huang, Maura Costello, Sabina Gude
seqWell, Inc. Beverly, MA USA

Introduction

The use of genomic prediction in agriculture has proven to be an effective strategy for accelerating breeding programs by enabling rapid and cost-efficient assessment of genetic potential. Genome-wide genotypes of sufficient density and accuracy can substitute for extensive phenotypic data collection, significantly shortening selection cycles. Several sequencing-based methods have been adopted to generate genotypes for these applications. Here we benchmarked the performance of AgriPrep™, a one-step, auto-normalizing high-throughput library preparation method optimized for low-pass whole genome sequencing (lpWGS), against traditional fragmentation and ligation library prep workflows for genomic selection using lpWGS in soybean, cattle and human samples. The three 8-plex libraries sequenced on the Illumina NovaSeq X Plus and downsampled to ~0.3×, ~0.6×, and ~1× coverage for comparison. To assess imputation quality, the coverage uniformity and genotype accuracy of human samples were analyzed. Our results demonstrate AgriPrep's suitability for routine, cost-effective lpWGS in agricultural genomics.

AgriPrep Library Preparation Kit Workflow

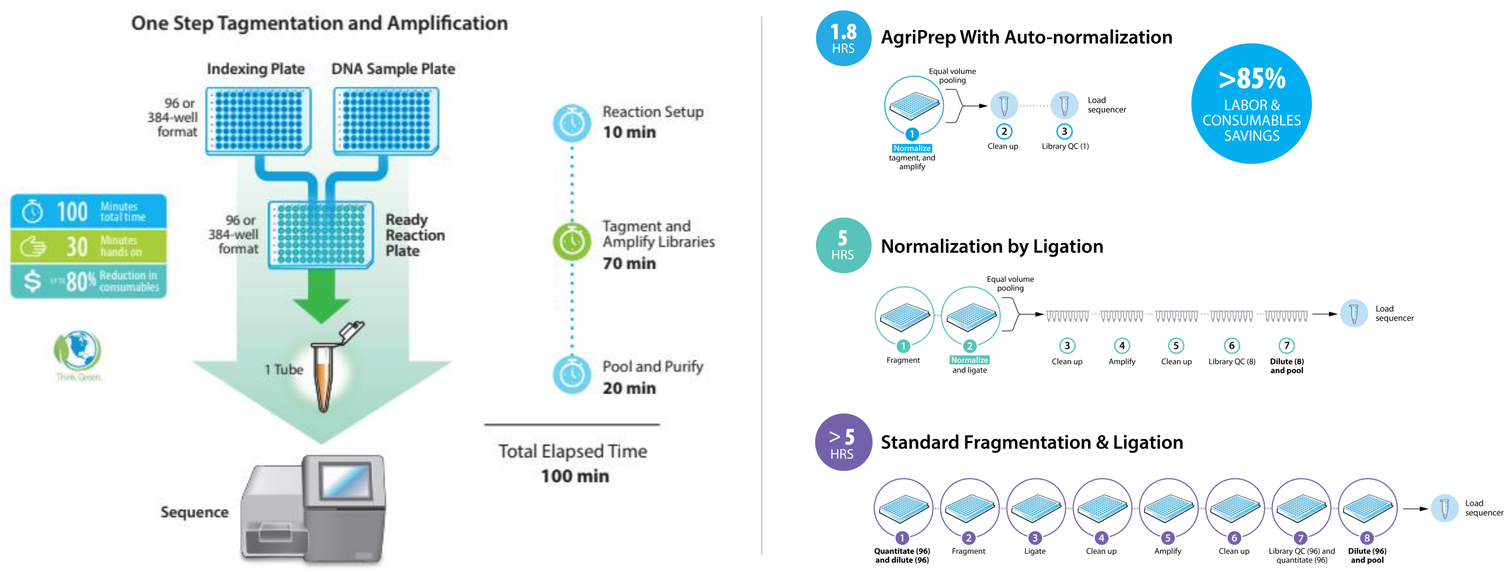

Figure 1. AgriPrep uses seqWell's high performance TnX™ transposase that was specifically engineered for NGS library preparation. The AgriPrep library prep kits utilize a proprietary mixture of enzymes to tag input DNA with indexed adapters and amplify libraries all in a single reaction. Samples are pooled volumetrically, purified, and converted into libraries to complete the 100-minute workflow, which includes 30 minutes of hands-on time. AgriPrep streamlines library preparation into a rapid, one-step workflow, reducing total time from >5 hours to ~1.8 hours while minimizing steps, consumables, and tip usage compared to traditional fragmentation and ligation methods.

Methods

- The AgriPrep library prep kit was used to processed soybean, cattle, and human samples using 30 ng input. Each sample was run in 8 replicates, creating three different pools of 8-plex.
- For comparison, a traditional enzymatic fragmentation and ligation library prep workflow and an enzymatic fragmentation, normalization-by-ligation library prep workflow were run side-by-side using the same input amount from each sample type in triplicate and quadruplicate (respectively) following manufacturer's protocol.
- The libraries prepared using three different library prep kits were sequenced on an Illumina NovaSeq X Plus (300 cycles).
- Sequencing data were demultiplexed, individually down-sampled to ~0.3×, ~0.6×, and ~1× coverage, and aligned via Picard Tools, before variant calling and imputation.
- The open-source GLIMPSE pipeline (v2.0.0) performed imputation using default settings for the human samples on chromosome 1.

Table 1. Summary of plant samples assessed in the study.

Samples	Genome Size (Gb)	Vendor/Catalog Number
Soybean *(Glycine max)*	1.1	Zyagen : PLG-1044
Cattle *(Bos taurus)*	2.7	Zyagen: GB-110F
Human *(Homo sapiens)*	3.1	Coriell ID: NA12878 (NIST ID: HG001)

Sequencing Quality Metrics

Table 2. Sequencing performance metrics summary of libraries prepared using three different library prep kits on the NovaSeq X Plus post down-sampling to 5.6M, 16M, and 18M random total reads (approximately 1× coverage for soybean, cattle, and human samples respectively).

Sample Type	Library Prep Method	# of Samples	Mean Coverage	Mean Insert Size	Duplication Rate	Estimated Library Size
Soybean *(Glycine max)*	AgriPrep	8	**1.16x**	373	9.9%	85.7 M
	Traditional Frag/Lig	3	1.12x	275	8.5%	90.4 M
	Normalization by Ligation	4	0.96x	221	5.7%	48.3 M
Cattle *(Bos taurus)*	AgriPrep	8	**1.11x**	426	10.3%	**240.5 M**
	Traditional Frag/Lig	3	1.10x	332	11.4%	214.8 M
	Normalization by Ligation	4	0.88x	229	10.0%	77.7 M
Human *(Homo sapiens)*	AgriPrep	8	**1.28x**	432	11.0%	**285.7 M**
	Traditional Frag/Lig	3	1.21x	331	12.8%	212.1 M
	Normalization by Ligation	4	0.95x	232	12.0%	67.4 M

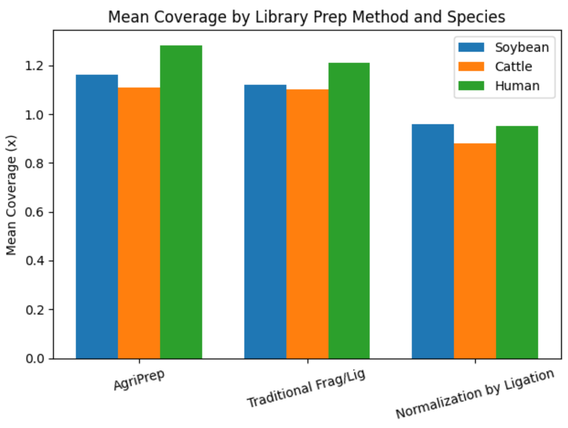

Figure 2. AgriPrep yields consistently higher mean sequencing coverage across species. Mean coverage across three species (Soybean, Cattle, and Human) is consistently higher with AgriPrep compared to traditional fragmentation/ ligation and normalization-by-ligation library prep methods. **For an equivalent number of read pairs, AgriPrep yields consistently higher mean sequencing coverage across species**, demonstrating improved usable sequencing depth with AgriPrep across diverse genomes.

Imputation Performance from Various Depths of Coverage

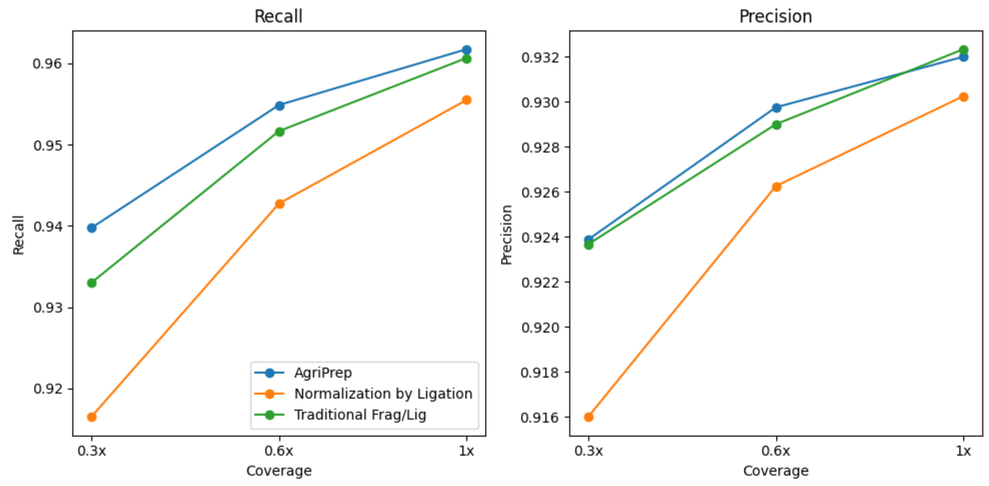

Figure 3. Recall (left) and precision (right) are shown as a function of sequencing coverage for each library preparation method. AgriPrep exhibits comparable precision and generally higher mean recall across coverage levels relative to comparator methods, with differences most visible at lower coverage.

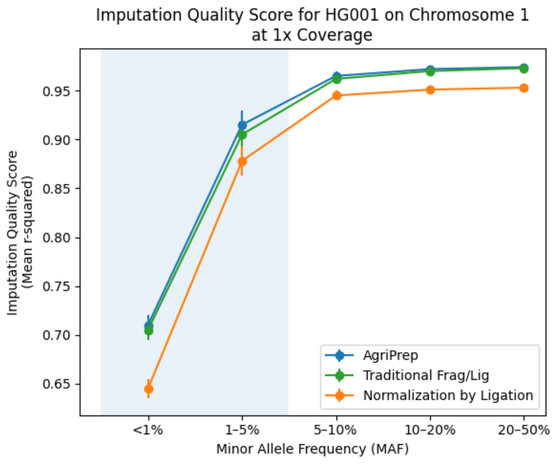

Figure 4. Mean imputation accuracy (r^2) is shown as a function of minor allele frequency (MAF) for HG001 on chromosome 1. Points represent mean values per MAF bin for each library prep method. AgriPrep exhibits comparable or higher mean performance across MAF bins relative to traditional fragmentation/ ligation and normalization-by-ligation library prep methods. The shaded region highlights low-frequency variants (<5% MAF), where greater variability is observed between the different library prep methods.

Summary and Conclusions

- AgriPrep library prep demonstrates superior performance across multiple sample types, delivering higher mean coverage, lower duplication rates, and significantly larger estimated library sizes compared to traditional fragmentation/ ligation and normalization-by-ligation library prep methods (Table 2, Figure 2). These improvements indicate enhanced library complexity and sequencing efficiency, enabling more effective use of sequencing reads and improved data quality across plant, animal, and human genomes.
- AgriPrep library prep shows comparable or higher mean performance across imputation accuracy and variant detection sensitivity under low-frequency and low-coverage conditions while maintaining comparable precision (Figure 3 and 4).

Download a Copy